

FIG. 1

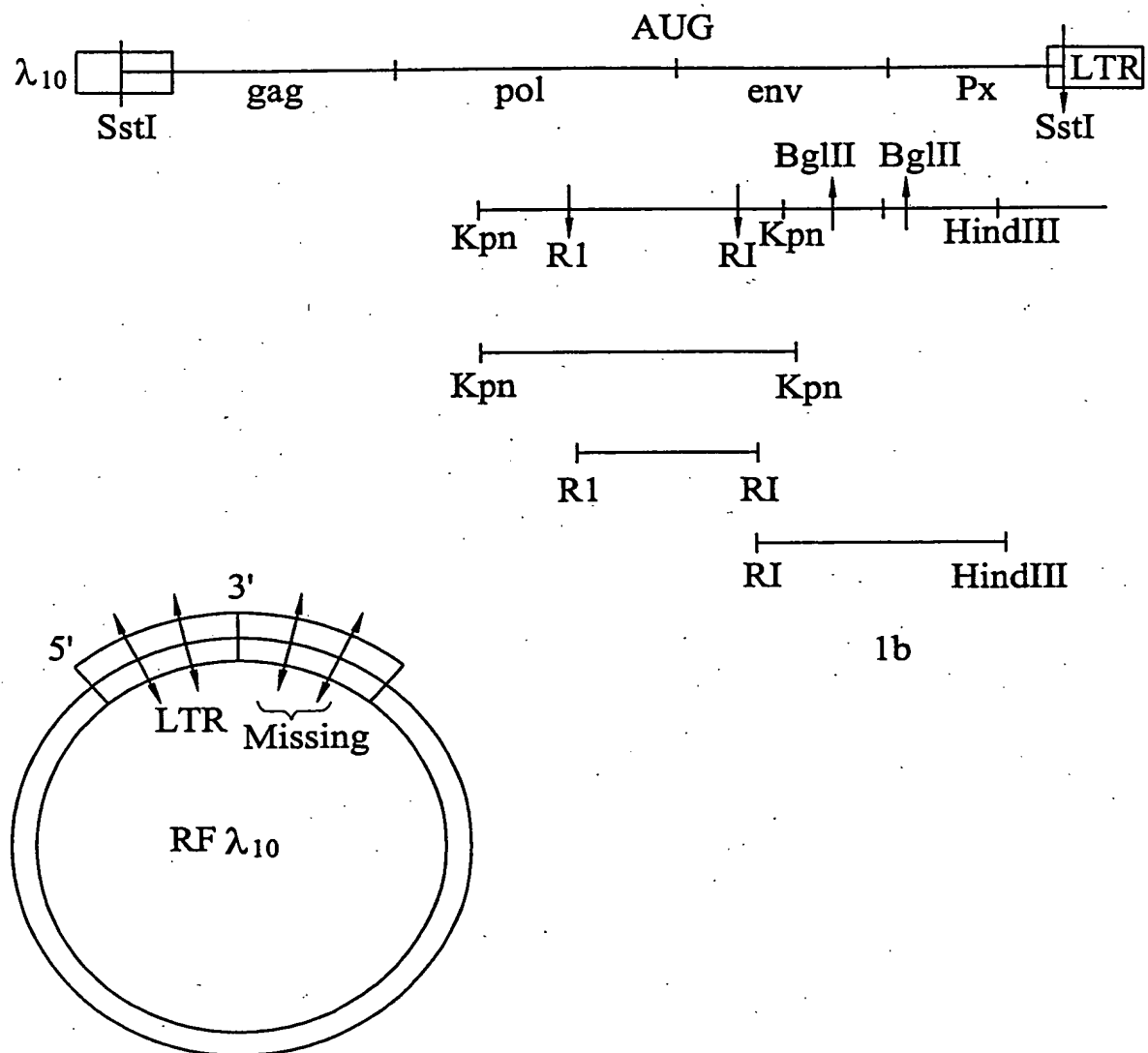
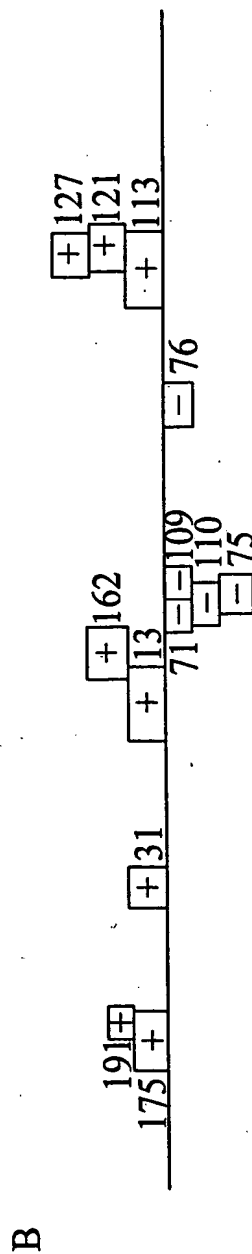
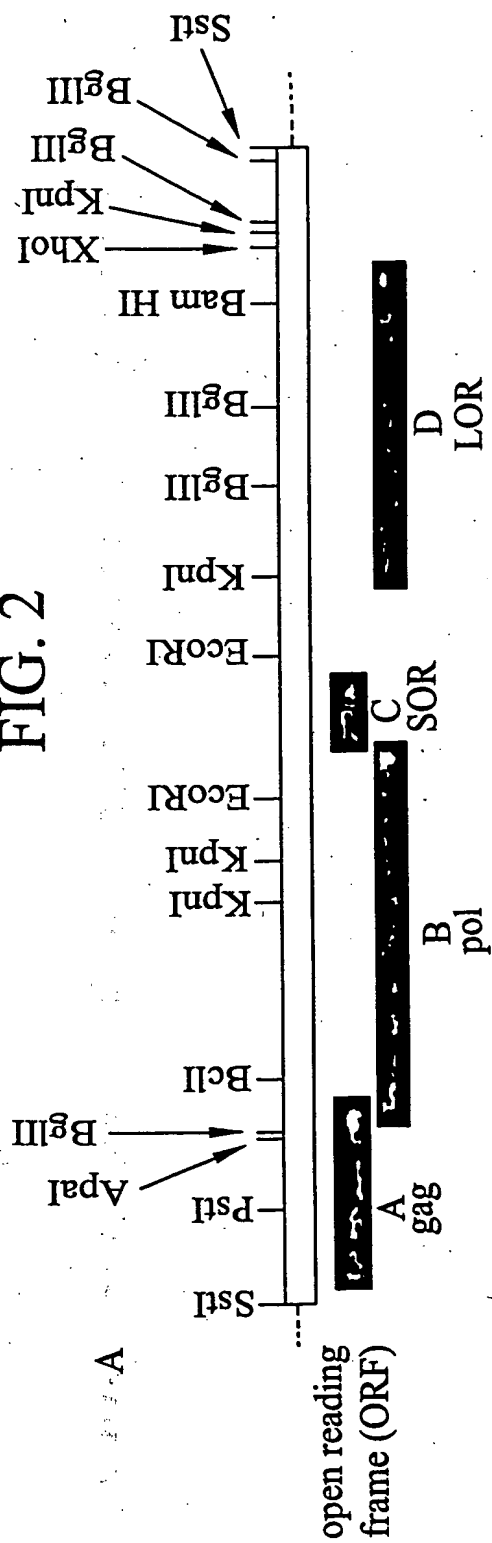


FIG. 2



[illegible]

FIG. 3 (Continued)

BH10 BH5	GAGCTCTCTCGACGAGGACTCGGCTTGCTGAAGCGGCACGGCAAGAGCGGCGGCGGCGACTGGTGAGTACG -----	296
BH10 BH5	CCAAAATTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGTGGGAGCGTCAGTATTAAAGCGGGGAGAATT ----- MetGlyAlaArqAlaSerValLeuSerGlyGlyGluLeu -----	371
BH10 BH5	AGATCGATGGGAAAAATTCCGTTAAGGCCAGGGGAAAGAAAAATATAATTAAACATATAGTATGGGCAAG A pArqTrpGluLysIleArqLeuArqProGlyGlyLysLysTyrLysLeuLysHisIleValTrpAlaSer -----	446
BH10 BH5	CAGGAGCTAGAACGATTCCGAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTOTAGACAAATACTGGGACA ArqGluLeuGluArqPheAlaValAsnProGlyLeuLeuGluThrSerGluGlyCysArqGlnIleLeuGlyGln -----	521
BH10 BH5	GCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCCTCTATTGT LeuGlnProSerLeuGlnThrGlySerGluGluLeuArqSerLeuTyrAsnThrValAlaThrLeuTyrCysVal ----- Hind III -----	596
BH10 BH5	GCATCAAAGGATAGAGATAAAGACACCAAGGAGCTTTAGACAAGATAGAGGAGCAAAACAAGTAAGAA HisGlnArqIleGluIleLysAspThrLysGluAlaLeuAspLysIleGluGluGluGlnAsnLysSerLysLys -----	671

13

38

63

88

113

FIG. 3 (Continued)

BH10	AAAAGCAGCAAGCAGCTGACACAGGACACAGCAGTCAGGTACGCCAAAATTACCTATAGTCAGAACAT	746	
BH5	LysAlaGlnGlnAlaAlaAlaAspThrGlyHisSerSerGlnValSerGlnAsnTyrProIleValGlnAsnIle	138	
	-----G-----		
BH10	CCAGGGGCAATGGTACATCAGGCCCATATCACCTAGAACTTTAAATGCATGGGTAAAGTAGTAGAAGAGAGGC	821	
BH5	GlnGlyGlnMetValHisGlnAlaIleSerProArqThrLeuAsnAlaTrpVallysValValGluGluLysAla	163	
	-----G-----		
BH10	TTTCAGCCCAAGTAATACCCATGTTTTTCAGCATTTATCAGAAGGAGCCACCCACCAAGATTTAAACACCCATGCT	896	
BH5	PheSerProGluValIleProMetPheSerAlaIleuSerGluGlyAlaThrProGlnAspLeuAsnThrMetLeu	188	
	-----G-----		
BH10	AAACACAGTGGGGGACATCAAGCAGCCATGCAAAATGTTAAAGAGAGACCATCAATGAGGAAGCTGCAGAAATGGGA	971	
BH5	AsnThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAlaAlaGluTrpAsp	213	
	-----G-----		
BH10	TAGAGTACATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGG	1046	
BH5	ArqValHisProValHisAlaGlyProIleAlaProGlyGlnMetArqGluProArqGlySerAspIleAlaGly	238	
	-----G-----		
BH10	AACTACTAGTACCCTTCAGGAACAAATAGGATGGATGACAAATAATCCACCTATCCCAGTAGGAGAAATTTATAA	1121	
BH5	ThrThrSerThrLeuGlnGluGlnIleGlyTrpMetThrAsnAsnProProIleProValGlyGluIleTyrLys	263	
	-----G-----		
BH10	AAGATGGATAATCCTGGGATTAAATAAGATGTAAGAAATGATAGCCCTACCAGCATTTCTGGACATTAAGACAAGG	1196	
BH5	ArqTrpIlelleuGlyLeuAsnLysIleValArqMetTyrSerProThrSerIleleuAspIleArqGlnGly	288	
	-----G-----		

Pvu II

GAG p17 ---|--- GAG p26

Aha III

Aha III

Pst I

FIG. 3 (Continued)

BH10	ACCAAAGAACCTTTAGAGACTATGTAGACCGGTCTCTATAAACTCTAAGAGCCGAGCAAGCTTCACAGGAGGT	1271	Hind III
BH5	ProLysGluProPheArqAspTyrValAspArqPheTyrLysThrLeuArqAlaGluGlnAlaSerGlnGluVal	313	
	-----G-----C-----A-----		
BH10	AAAAATTGGATGACAGAAACCTTGTGGTCCAAAATGCCAACCAGATTGTAGACTATTTTAAAGCATTGGG	1346	Aha III
BH5	LysAsnTrpMetThrGluThrLeuValGlnAsnAlaAsnProAspCysLysThrIleLeuLysAlaLeuGly	338	

BH10	ACCAGCGGTACACTAGAAGAAATGATGACAGCATGTCAGGGAGTAGGAGGCCCGCCATAAGCAAGAGTTT	1421	
BH5	ProAlaAlaThrLeuGluGluMetMetThrAlaCysGlnGlyValGlyGlyProGlyHisLysAlaArqValLeu	363	

BH10	GGCTGAAGCAATGAGCCCAAGTAACAATACAGCTACCATTAATGATGCAGAGAGGCAATTTTAGGAACCAAGAAA	1496	
BH5	AlaGluAlaMetSerGlnValThrAsnThrAlaThrIleMetMetGlnArqGlyAsnPheArqAsnGlnArqLys	388	
	-----T-----A-----A-----SerThr		
BH10	GATGTTAAGTGTTCATTGTGGCAAGAAGGACACAGCCAGAAATTGCAGGGCCCTAGGAAAGGGCTG	1571	
	MetValLysCysPheAsnGlyLysGlyLysGlyHisThrAlaArqAsnGlyArqAlaProArqLysLysGlyCys	413	
	A---T-----A-----T-----A-----Lys		
	-----Ile-----Arg		
	-----Direct Repeat-----POL-----Bql II		
BH10	TTGGAAATGTGGAAAGGACACCAAAATGAAGATTGTACTGAGAGACAGGCTAATTTTATAGGAAGATCTG	1646	
	TrpLysCysGlyLysGlyHisGlnMetLysAspCysThrGluArqGlnAlaAsnPheLeuGlyLysIleTrp	438	
	-----PhePheArqGluAspLeu-----	6	
BH5	-----		
BH10	GCCTTCCTACAAGGAAGGCCAGGGAATTTTCTTCAGAGCAGACAGCCCAACAGCCCACTTCTTCAGAG	1721	Direct
	ProSerTyrLysGlyArqProGlyAsnPheLeuGlnSerArqProGluProThrAlaProProPheLeuGlnSer	463	
	AlaPheLeuGlnGlyLysAlaArqGluPheSerSerGluGlnThrArqAlaAsnSerProThrIleSerSerGlu	31	
BH5	-----		

FIG. 3 (Continued)

BH10	Repeat	1796	488
BH5	CAGACCCAGAGCCCAACAGCCACAGAGAGCTTCAGGCTCGGGTAGAGACAACAACCTCCCTCAGAAAGCA ArqPr GluProThrAlaProProGluGluSerPheArqSerGlyValGluThrThrThrProProGlnLysGln GlnThrArqAlaAsnSerProThrArqArqGluLeuGlnValTrpGlyArqAsnAsnSerProSerGluAla ----- Ser Leu GAG p15		56
BH10	GGAGCCGATAGACAAGGAACGTATCCTTTAACTTCCCTCAGATCACTCTTTGCCAACGACCCCTCGTCACAATA GluProIleAspLysGluLeuTyrProLeuThrSerLeuArqSerLeuPheGlyAsnAspProSerSerGln GlyAlaAspArqGlnGlyThrValSerPheAsnPheProGlnIleThrLeuTrpGlnArqProLeuValThrIle -----	1871	512
BH5	-----		81
BH10	AAGATAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTG LysIleGlyGlyGlnLeuLysGluAlaIleLeuAspThrGlyAlaAspAspThrValLeuGluGluMetSerLeu -----	1946	106
BH5	-----		
BH10	CCAGGAAGATGGAAACCAAAATGATAGGGGAATTGGAGGTTTTATCAAGTAAGACAGTATGATCAGATACTC Pr GlyArqTrpLysProLysMetIleGlyGlyIleGlyGlyPheIleLysValArqGlnTyrAspGlnIleLeu -----	2021	131
BH5	-----		
BH10	ATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAAT II GluIleCysGlyHisLysAlaIleGlyThrValLeuValGlyProThrProValAsnIleIleGlyArqAsn -----	2096	156
BH5	----- Aha III		
BH10	CTGTTGACTCAGATTGGTTGCACCTTTAAATTTCCCATTTAGCCCTATTGAGACTGTACCAGTAAATTAAGCCA LeuLeuThrGlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrValProValLysLeuLysPro ----- A	2171	181
BH5	-----		

FIG. 3 (Continued)

BH10	2246	206	
BH5			
BH10	2321	231	
BH5			
BH10	2396	256	
BH5			
BH10	2471	281	
BH5			
BH10	2546	306	
BH5			
BH10	2621	331	
BH5			
BH10	2696	356	
BH5			

GGAATGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAGAAGAAAAATAAAGCATTAGTAGAAAATTTGTACA
 GlyMetAspGlyProLysValLysGlnTrpProLeuThrGluGluLysIleLysAlaLeuValGluIleCysThr

 GAAATGGAAAAGGAGAAATTTCAAAAATTGGCCCTGAGAAATCCATACAAATACTCCAGTATTTGCCATAAAG
 GluMetGluLysGluGlyLysIleSerLysIleGlyProGluAsnProTyrAsnThrProValPheAlaIleLys

 AAAAAGACAGTACTAAATGGAGAAAATTAGTAGATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAA
 LysLysAspSerThrLysTrpArqLysLeuValAspPheArqGluLeuAsnLysArqThrGlnAspPheTrpGlu

 GTTCAATTAGGAATACCACATCCCGCAGGGTTAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCA
 ValGlnLeuGlyIleProHisProAlaGlyLeuLysLysLysSerValThrValLeuAspValGlyAspAla

 TATTTTTCAGTTCCTTAGATGAAGACITTCAGGAAGTATACCTGCAATTTACCATACCTAGTATAAACAATGAGACA
 TyrPheSerValProLeuAspGluAspPheArqLysTrpThrAlaPheThrIleProSerIleAsnAsnGluThr

 CCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCATG
 ProGlyIleArqTrpGlnTyrAsnValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSerSerMet

 ACAAATCTTAGAGCCTTTTAAAAAACAAAAATCCAGACATAGTTATCTATCAATACATGGAATGATTTGTATGTA
 ThrLysIleLeuGluProPheLysLysGlnAsnProAspIleValIleLysGlnTyrMetAspLysLeuLysVal

Arg

Asp III

Arg

FIG. 3 (Continued)

BH10	GGATCTGACTTAGAAATAGGGCAGCATAGAACAAATAAGAGGAGCTGAGACAACATCTGTTGAGGTGGGACTT	2771	381
BH5	GlySerAspLeuGluIleGlyGlnHisArqThrLysIleGluGluLeuArqGlnHisLeuEuArqTrpGlyLeu		
	-----Phe		
BH10	ACCACACCAGACAAACATCAGAAAGAACCTCCATTCTTGGTGGTTATGAACCTCCATCCTGATAAATGG	2846	406
BH5	ThrThrProAspLysLysHisGlnLysGluProProPheLeuTrpMetGlyTyrGluLeuHisProAspLysTrp		
BH10	ACAGTACAGCCTATAGTGTGCCAGAAAGACAGCTGGACTGTCAATGACATACAGAAGTTAGTGGGAAATG	2921	431
BH5	ThrValGlnProIleValLeuProGluLysAspSerTrpThrValAsnAspIleGlnLysLeuValGlyLysLeu		
	-----A		
	-----Ile		
BH10	AATTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTATGTAACTCCTTAGAGGAACCAAGCACTA	2996	456
BH5	AsnTrpAlaSerGlnIleTyrProGlyIleLysValArqGlnLeuCysLysLeuEuArqGlyThrLysAlaLeu		
BH10	ACAGAAGTAATACCACCTAACAGAAGAGCAGAGCTAGAACTGGCAGAAACAGAGAGATTCTAAAGAACCAGTA	3071	481
BH5	ThrGluValIleProLeuThrGluGluAlaGluLeuGluLeuAlaGluAsnArqGluIleLeuLysGluProVal		
BH10	CATGGAGTGTATTATGACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCCAATGGACATAT	3146	506
BH5	HisGlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGlnGlyGlnGlnTrpThrTyr		
	-----Aha III		
BH10	CAAAATTTACAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAAGAAATGAGGGGTGCCACACATAATGAT	3221	531
BH5	GlnIleTyrGlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArqMetArqGlyAlaHisThrAsnAsp		
	-----Aha III		
BH10	GTAAACAAATTAACAGAGGCAGTGCAAAAAATAACCACAGAAAGCATAGTAATATGGGAAAGACTCCTAAATTT	3296	556
BH5	ValLysGlnLeuThrGluAlaValGlnLysIleThrThrGluSerIleValIleTrpGlyLysThrProLysPhe		

FIG. 3 (Continued)

BH10	AACTACCCATACAAAGGAAACATGGGAAACATGGTGACAGAGTATTGGCAAGCCACCTGGATTCTCCTGAGTGG	3371	581
BH5	LysLeuProIlaGlnLysGluThrTrpGluThrTrpTrpThrGluTyrTrpGlnAlaThrTrpIleProGluTrp		
	-----A-----		
BH10	GAGTTTGTTAATACCCCTCCTTTAGTGAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACC	3446	606
BH5	GluPheValAsnThrProProLeuValLysLeuTrpTyrGlnLeuGluLysGluProIleValGlyAlaGluThr		
	-----Kpn I-----		
BH10	TTCTATGTAGATGGGGCAGCTAACAGGGAGACTAAATTAGGAAAGCAGGATATGTTACTAACAAAGGAAGACAA	3521	631
BH5	PheTyrValAspGlyAlaAlaAsnArqGluThrLysLeuGlyLysAlaGlyTyrValThrAsnLysGlyArqGln		
	-----G-----Ser-----Arg-----		
BH10	AAGTTGTCCCCCTAACTAACACACAACAATCAGAAACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCA	3596	656
BH5	LysValValProLeuThrAsnThrThrAsnGlnLysThrGluLeuGlnAlaIleTyrLeuAlaLeuGlnAspSer		
	---A---C---G---A---C---Asn-----		
	Thr His		
BH10	GGATTAGAAGTAAACATAGTAACAGACTCACAAATATGCATTAGGAATCATTCAAGCACACACCAGATAAAGTGAA	3671	681
BH5	GlyL uGluValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAlaGlnProAspLysSerGlu		
	-----I-----		
BH10	TCAGAGTTAGTCAATCAAATAATAGAGCAGTTAATAAAAGGAAAGGTCTATCTGGCATGGGTACCAGCACAC	3746	706
BH5	SerGluLeuValAsnGlnIleIleGluGlnLeuIleLysLysGluLysValTyrLeuAlaTrpValProAlaHis		
	-----Kpn I-----		
BH10	AAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGTGGAATCAGGAAATACTATTTTAGATGGA	3821	731
BH5	LysGlyIleGlyGlyAsnGluGlnValAspLysLeuValSerAlaGlyIleArqLysIleLeuPheLeuAspGly		

FIG. 3 (Continued)

BH10	ATAGATAAGGCCCAAGATGAACATGAGAAATATCACAGTAAATTGGAGAGCAATGGCTAGTGTATTTTAACTGCCA	3896	756
BH5	IleAspLysAlaGlnAspGluHisGluLysTyrHisSerAsnTrpArqAlaMetAlaSerAspPheAsnLeuPro		
	-----A-----		
BH10	CCTGTAGTAGCAAAAGAAATAGTAGCCAGCTGTGATAAATOTCAGCTAAAGGAGAGCCATGCATGGACAAGTA	3971	781
BH5	Pr ValV lAlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGluAlaMetHisGlyGlnVal		
	-----Pvu II-----		
BH10	GACTGTAGTCCAGGAAATATGGCAACTAGATTGTACACATTTAGAAGGAAAGTTATCCTGCTAGCAGTTTCATGTA	4046	806
BH5	AspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeuGluGlyLysValIleLeuValAlaValHisVal		
	-----Aha III-----		
BH10	GCCAGTGGATATATAGAAGCAGAAAGTTATCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTTAAATTA	4121	831
BH5	AlaSerGlyTyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeuLysLeu		

BH10	GCAGGAAGATGGCCAOTAAAAACAATACATACAGACAATGGCAGCAATTTCCACCAGTGTACGGTTAAGGCCGCC	4196	856
BH5	AlaGlyArqTrpProValLysThrIleHisThrAspAsnGlySerAsnPheThrSerAlaThrValLysAlaAla		
	-----Eco RI-----		
BH10	TOTTTGGTGGCGGGAATCAAGCAGGAATTTGGAAATTCCTTACAATCCCCAAAGTCAAGCAGTAGTAGAATCTATG	4271	881
BH5	CysTrpTrpAlaGlyIleLysGlnGluPheGlyIleProTyrAsnProGlnSerGlnGlyValValGluSerMet		

BH10	AATAAGAAATTAAGAAATATAGGACAGGTAAGAGATCAGGCTGAACATCTTAAGACAGCAGTACAAAATGCCA	4346	906
BH5	AsnLysGluLeuLysLysIleIleGlyGlnValArqAspGlnAlaGluHisLeuLysThrAlaValGlnMetAla		

FIG. 3 (Continued)

BH10	931	4421	GTATTCATCCACAATTTTAAAGAAAAGGGGGATYGGGGGTACAGTGCAGGGGAAAGAAATAGTACATAATA
BH5			ValPheIleHisAsnPhelysArqlysGlyGlyIleGlyGlyTyrSerAlaGlyGluArqlleValAspIlelle
BH10	956	4496	GCAACAGACATACAAACTAAAGAAATTACAAAACAAATTACAAAATTCAAAAATTTTCGGGTTTTATTACAGGGAC
BH5			AlaThra pIleGlnThrLysGluLeuGlnLysGlnIleThrLysIleGlnAsnPheArqValTyrTyrArqAsp
BH10	981	4571	AGCAGAAATCCACTTTGGAAAGGACCAGCAAGCTCCTCTGGAAAGGTGAAGGGCAGTAGTAATACAAAGATAAT
BH5			SerArqAsnProleuTrpLysGlyProAlaLysLeuLeuTrpLysGlyGlyAlaValValIleGlnAspAsn
BH10	1006	4646	AGTGACATAAAGTAGTGCCCAAGAAAGCAAGATCATYAGGGATTATGGAAAACAGATGGCAGGTGATGAT
BH5	20		SerAspIleLysValValProArqArqlysAlaLysIleIleArqAspTyrGlyLysGlnMetAlaGlyAspAsp
			CysGlnGluGluLysGlnArqSerLeuGlyIleMetGluAsnArqTrpGlnValMetIle
BH10	1015	4721	TGTGTGGCAAGTAGACAGGATGAGGATTAGAACATGGAAAGTTAGTAAACACCATATGTATGTTTCAGGGAA
BH5	45		CysV lAlaSerArqGlnAspGluAsp
			ValTrpGlnValAspArqMetArqIleArqThrTrpLysSerLeuValLysHisMetTyrValSerGlyLys
			Arg
BH10	70	4796	AGCTAGGGGATGGTTTTATAGACATCACTATGAAAGCCCTCATCCAAGAATAAGTTCAGAAGTACACATCCCAC
BH5			AlaArqGlyTrpPheTyrArqHisTyrGluSerProHisProArqlleSerSerGluValHisIleProleu
BH10	95	4871	AGGGGATGCTAGATTGGTAATAACAACATATTGGGGTCTGCATACAGGAGAAAGAGACTGGCATTGGGTACAGG
BH5			GlyAspAlaArqLeuValIleThrThrTrpGlyLeuHisThrGlyGluArqAspTrpHisLeuGlyGlnGly

FIG. 3 (Continued)

BH10	AGTCTCCATAGAAATGGAGGAAAGAGATATAGCACACAAGTAGACCCCTGAACCTAGCAGACCAACTAATTCATCT	4946
BH5	ValSerIleGluTrpArqLysLysArqTyrSerThrGlnValAspProGluLeuAlaAspGlnLeuIleHisLeu	120
	-----G-----	
	Arg	
BH10	GTATTACTTTGACTGTTTTTCAGACTCTGCTATAGAAAGGCCCTTATTAGGACACATAGTTAGCCCTAGGTGCTGA	5021
BH5	TyrTyrPheAspCysPheSerAspSerAlaIleArqLysAlaLeuLeuGlyHisIleValSerProArqCysGlu	145
	-----T-----	
BH10	ATATCAAGCAGGACATAACAAGGTAGGATCTCTACAATACTTGGCAGTACGAGCATTAAATAACACCAAAAAGAT	5096
BH5	TyrGlnAlaGlyHisAsnLysValGlySerLeuClnTyrLeuAlaLeuAlaLeuIleThrProLysLysIle	170
	-----G-----	
	Val	
BH10	AAAGCCACCTTTGCCTAGTGTACGAAACTGACAGAGGATAGATGGAACAAGCCCAAGAGCAAGGCCACAG	5171
BH5	LysPr ProLeuProSerValThrLysLeuThrGluAspArqIrpAsnLysProGlnLysThrLysGlyHisArq	195
	-----A-----	
	SOR	
BH10	AGGAGCCACACAAATGAATGGACACTAGAGCTTTTAGAGGAGCTTAAGAATGAAGCTGTTAGACATTTTCCTAGG	5246
BH5	GlySerHisThrMetAsnGlyHis	203
	-----A-----	
BH10	ATTTGGCTCCATGGCTTAGGGCAACATATCTATGAACTTATGGGATACTTGGCAGGAGTGGAGGCCATAATA	5321
BH5	-----A-----	
	Eco RI	
BH10	AGAATTCTGCAACAACACTGCTGTTTATCCATTTTCAGAAATGGGTGTTCGACATAGCAGAATAGCGTTACTCGACA	5396
BH5	-----A-----	

FIG. 3 (Continued)

BH10 BH5	GAGGAGAGCAAGAAATGGAGCCAGTAGATCCTAGACTAGAGCCCTGGAAGCATCCAGGAAGTCAGCCTAAACTG -----	5471
BH10 BH5	CYTGTACCAATTGCTATTGTAAAAGTGTGCTTTCATTGCCAAGTTGTTTCATAACAAAGCCTTAGGCATCT -----C-----	5546
BH10 BH5 BH8	CCTATGGCAGGAAGAACGGGAGACAGCGACCGAAGACCCTCCTCAAGGCAGTCAGACTCATCAAGTTTCTCTATCAA ----- -----G----- --G---A--G-A-----	5621
BH10 BH8	AGCAGTAAGTAGTACATGTAAATGCAACCTATACAAATAGCAATAGTAGTACAT JTTAGTAGTAGCAATAATAAGCAA -----C-----T--C-A---T-GCC---C-----	5696
BH10 BH8	TAGTTCTGTGGTCCATAGTAATCATAGAATATAGGAAATATTAGACAAAGAAATAGACAGGTTAATTGATA ----- -----ENV-LOR-----	5771
BH10 BH8	GACTAATAGAAAGAGCAGAGACAGTGGCAATGAGAGTGAAGGAGAAATATCAGCAGCTTGTGGAGATGGGGGTGG -----LysGluGlnLysThrValAlaMetArqValLysGluLysIyrGlnHisLeuTrpArqTrpGlyTrp-----	5846
BH10 BH8	AGATGGGGCACCATGCTCCTTGGGATGTTGATGATCTGTAGTGTGCTACAGAAATAATTGTGGTCCACAGTCTATTAT ArqTrpGlyThrMetLeuLeuGlyMetLeuMetIleCysSerAlaThrOluLysLeuTrpValThrValIyrIyr ----- -----Phe-----	5921
BH10 BH8	GGGTACCTGTGTGGAGGAAGCAACCACCACCTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAGAGGTA Kpn I GlyValProValTrpLysGluAlaThrThrLeuPheCysAlaSerAspAlaLysAlaIyrAspThrGluVal -----	5996

22

47

72

FIG. 3 (Continued)

BH10	CATAATGTTGGCCACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTAGTATTGGTAAATGTGACA	6071
BH8	HisAsnValTrpAlaThrHisAlaCysValProThrAspProAsnProGlnGluValValLeuValAsnValThr	97
BH10	GAAATTTTAAACATGTGGAAAAATGACATGGTAGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAGC	6146
BH8	GluAsnPheAsnMetTrpLysAsnAspMetValGluGlnMetHisGluAspIleIleSerLeuTrpAspGlnSer	122
BH10	CTAAGCCATGTGTAAATTAACCCCACTCTGTGTAGTTTAAAGTGCACCTGATTGGAAGATGATACTAATACC	6221
BH8	LeuLysProCysValLysLeuThrProLeuCysValSerLeuLysCysThrAspLeuLysAsnAspThrAsnThr	147
BH10	AATAGTAGCGGAGAAATGATAATGGAGAAAGGAGAGATAAAAAACCTGCTCTTTCATATATCAGCACAAAGCATA	6296
BH8	AsnSerSerGlyArgMetIleMetGluLysGlyGluIleLysAsnCysSerPheAsnIleSerThrSerIle	172
	-----A-----Lys	
BH10	AGAGGTAAGGTGCAGAAAGAATATGCATTTTATAACTTGATATAATACCAATAGATAATGATACTACCAGC	6371
BH8	ArgGlyLysValGlnLysGluTyrAlaPhePheTyrLysLeuAspIleIleProIleAspAsnAspThrThrSer	197
BH10	TATACGTTGACAAGTTGTAAACACCTCAGTCATTACACAGGCTGTCCAAAGGTATCCTTTGAGCCCAATTCCCATA	6446
BH8	TyrThrLeuThrSerCysAsnThrSerValIleThrGlnAlaCysProLysValSerPheGluProIleProIle	222
BH10	CATTATTGTGCCCGGCTGGTTTTCGGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACA	6521
BH8	HisTyrCysAlaProAlaGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGlyThrGlyProCysThr	247
BH10	AATGTCAGCACAGTACAAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTAAATGGCAGTCTG	6596
BH8	AsnValSerThrValGlnCysThrHisGlyIleArgProValValSerThrGlnLeuLeuAsnGlySerLeu	272

FIG. 3 (Continued)

[illegible]

FIG. 3 (Continued)

BH10	GGACAAATAGATGTTCAATAATTACAGGGCTGCTATTAAACAAGAGATGGTGGTAATAGCAACAATGAGTCC	7196	472
BH8	GlyGlnIleArqCysSerSerAsnIleThrGlyLeuLeuThrArqAspGlyGlyAsnSerAsnAsnGluSer		
BH10	Bq1 II GAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAATATAAAGTAGTAAAA	7271	497
BH8	GluIlePheArqProGlyGlyGlyAspMetArqAspAsnTrpArqSerGluLeuTyrLysTyrLysValValLys		
BH10	ATTGAACCATYAGGAGTAGCACCCACCAAGGCAAGAGAGAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATA	7346	522
BH8	IleGluProLeuGlyValAlaProThrLysAlaLysArqArqValValGlnArqGluLysArqAlaValGlyIle		
BH10	GGAGCTTTGTTCCCTTGGGTTCTTGGGAGCAGCAGGAAGCACATATGGCGGCAGCOTCAATGACGGTGACGGTACAG	7421	547
BH8	GlyAlaLeuPheLeuGlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaAlaSerMetThrLeuThrValGln		
BH10	GCCAGACAAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAATTTGCTGAGGGCTATTGAGGGCGCAACAGCATCTG	7496	572
BH8	AlaArqGlnLeuLeuSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArqAlaIleGluAlaGlnGlnHisLeu		
	-----GC-----Gly		
BH10	TTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAAATCCTGGCTGTGGAAAGATACCTAAAGGATCAA	7571	597
BH8	LeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArqIleLeuAlaValGluArqTyrLeuLysAspGln		
BH10	CAGCTCCTGGGGATTGGGGTTGCTCTGGAAAACACTATTTCACCACCTGCTGTGCCCTTGGAAIGCTAGTTGGAGT	7646	622
BH8	GlnLeuLeuGlyIleTrpGlyCysSerGlyLysLeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSer		

FIG. 3 (Continued)

BH10	647	7721	7721
BH8			
BH10	672	7796	
BH8			
BH10	697	7871	
BH8			
BH10	722	7946	
BH8			
BH10	747	8021	
BH8			
BH10	772	8096	
BH8			
BH10	797	8171	
BH8			

M Hind III
 AATAAATCTCTGGAACAGATTGGGAATAACATGACCTGGATGGAGTGGACAGAGAAATTAAACAATTACACAAGC
 AsnLysSerLeuGluGlnIleTrpAsnAsnMetThrTrpMetGluTrpAspArqGluIleAsnAsnTyrThrSer

 TTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAGAATGAACAAGAATTATTGGAAATTAGATAAA
 LeuIleHisSerLeuIleGluGluSerGlnAsnGlnGlnGlnLysAsnGluGlnGluLeuGluLeuAspLys

 TGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAGTAGGA
 TrpAlaSerLeuTrpAsnTrpPheAsnIleThrAsnTrpLeuTrpTyrIleLysLeuPheIleMetIleValGly

 GGCTTGGTAGGTTTAAGAATAGTTTTTGTCTGTACTTTCGTAGTGAATAGAGTTAGCAGGGATATTACACCATTA
 GlyLeuValGlyLeuArqIleValPheAlaValLeuSerValValAsnArqValArqGlnGlyTyrSerProLeu

 TCGTTTCAGACCCACCTCCCAATCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAGGTGGAGAGAGA
 SerPheGlnThrHisLeuProIleProArqGlyProAspArqProGluGlyIleGluGluGluGlyGlyGluArq

 GACAGAGACAGATCCATTAGTGAACGGATCCTTAGCAGCTTATCTGGGACGATCTGGGAGCCTGTGCCTC
 AspArqAspArqSerIleArqLeuValAsnGlySerLeuAlaLeuIleTrpAspAspLeuArqSerLeuCysLeu

 TTCAGCTACCACCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTTCTGGGACGCAGGGGTGG
 PheSerTyrHisArqLeuArqAspLeuLeuIleValThrArqIleValGluLeuGlyArqArqGlyTrp

Ile

Asn

FIG. 3 (Continued)

BH10	GAAGCCCTCAAATATTGGTGGAAATCTCCTACAGTATTGGAGTCAGGAGCTAAAGAATAGTGCTGTTAGCTTGCTC	8246
BH8	GlualaleuLysTyrTrpTrpAsnLeuleuGlnTyrTrpSerGlnGluLeuLysAsnSerAlaValSerLeuleu	822
	-----A-----A-----Asn	
BH10	AATGCCACAGCTATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTATAGAGCTATT	8321
BH8	AsnAlaThrAlaIleAlaValAlaGluGlyThrAspArgValIleGluValValGlnGlyAlaTyrArgAlaIle	847
	-----C-----C-----Ala	
BH10	CGCCACATACCTAGAAGAATAAGACAGGGCTTGGAAAGGATTTTCTCTATAAGATGGGTGGCAAGTGGTCAAAAAG	8396
BH8	ArgHisIleProArgArgIleArgGlnGlyLeuGluArgIleLeuleu	863
	-----T-----T-----	
BH10	TAGTGTGGTTGGATGGCCTGCTGTAAAGGAAAGAATGAGACGAGCTGAGCCAGCAGCAGATGGGTGGGAGCAGC	8471
BH8	-----T-----T-----	
BH10	ATCTCGAGACCTAGAAAAACATGGAGCAATCACAAAGTAGCAACACAGCAGCTAACAAATGCTGATTGTGCCCTGGCT	8546
BH8	-----T-----T-----C-----C-----	
	-----Xh I-----Kpn I-----	
BH10	AGAAGCACAGAGGAGGAGGTGGGTTTTCCAGTCACACCTCAGGTACCTTTAAGACCAATGACTTACAAGGC	8621
BH8	-----PvuII BqI II-----Aha III-----U3	
	-----Polypurine Tract IR-----	
BH10	AGCTGTAGATCTTAGCCACTTTTAAAGAAAGGGGAGCTGGAAGGGCTAATTCACCTCCCAACGAAGACAAGA	8696
BH8	-----C-----C-----	
BH10	TATCCTTGATCTGTGGATCTACCCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCGGGCCAGGGAT	8771
BH8	-----C-----G-----AG-----	
	-----Bam HI-----	

BH10 BH8	CAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCAGAGAAGTTAGAAGAACCCACAA -----A-----T-----	8846
BH10 BH8	AGGAGAGAACACCAGCTTGTACACCCCTGTGAGCCYGCATGGAA TGGATGACCCGGAGAGAGAAGTGTAGAGTG -----T-----	8921
BH10 BH8	GAGGTTGACAGCCGCTAGCATTTCAATCACATGGCCCCGAGAGCTGCATCCGGAGTACTTCAAGAAGCTGCTGACA -----T-----	8996
BH10 BH8	TCGAGCTTGCTACAAGGACTTTCGGCTGGGACTTTCAGGGAGGCGTGGCTGGCGGGACTGGGGAGTGGCG -----	9071
BH10 BH8	AGCCCTCAGATCCTGCATATAAGCAGCTGCTTTTGGCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCT ----- Sst I R GGGAGCTC ----- Pvu II U3-----R Bgl II	9146
BH10 BH8		9154
HXB2	TCTGGCTAGCTAGGGAACCCACGCTTAAGCCCTCAATAAAGCTTGCCCTGAGTCTTCA ----- Hind III Poly(A) Sig. R----- U5 AGTAGTGTGTGCCCCGCTCTGTGTGTGACTCTGGTAACTAGAGATCCCTCAGA	9213
HXB2	CCCTTTTTCAGTCAGTGTGGAAATCTCTAGCA ----- IR	

FIG. 4

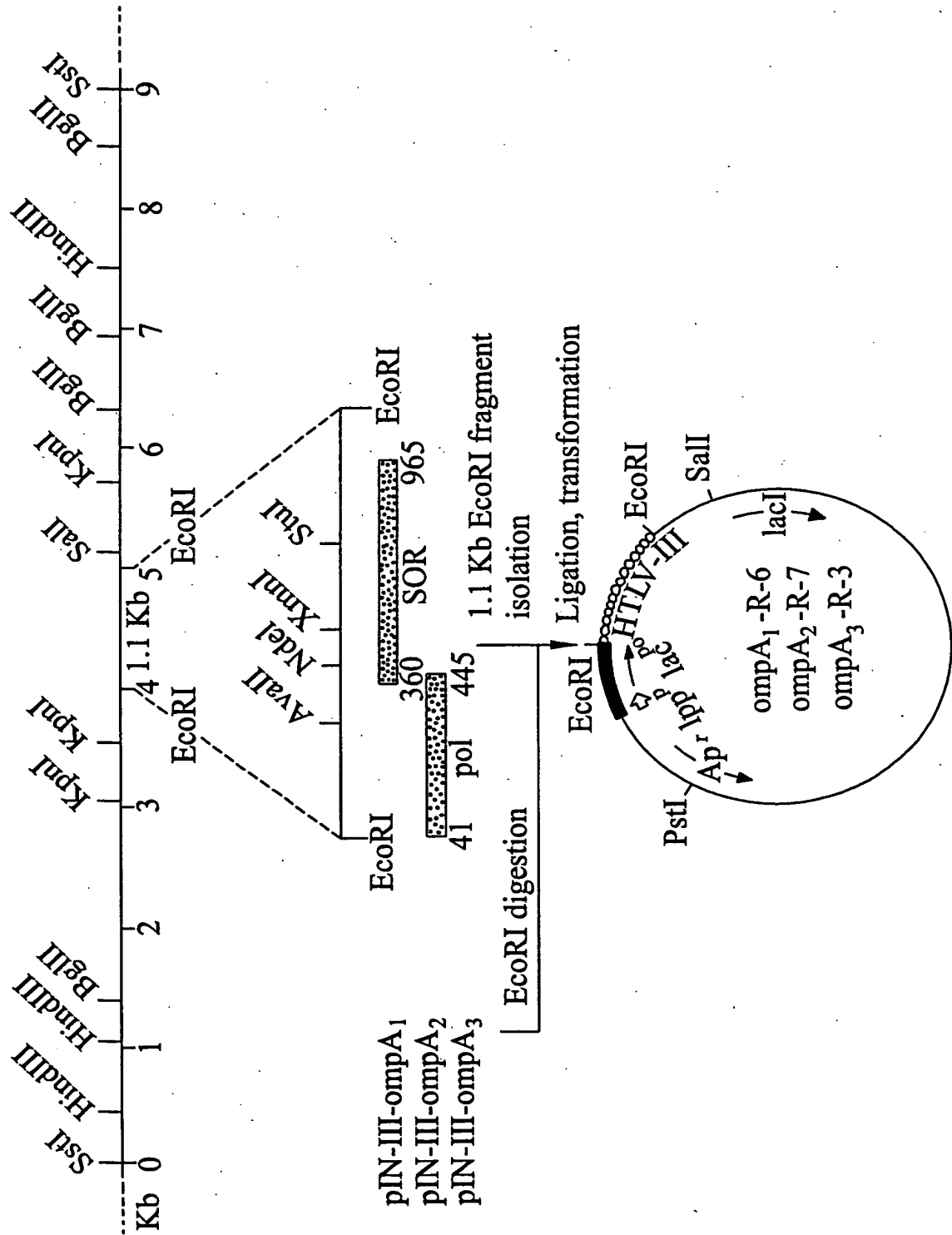


FIG. 5

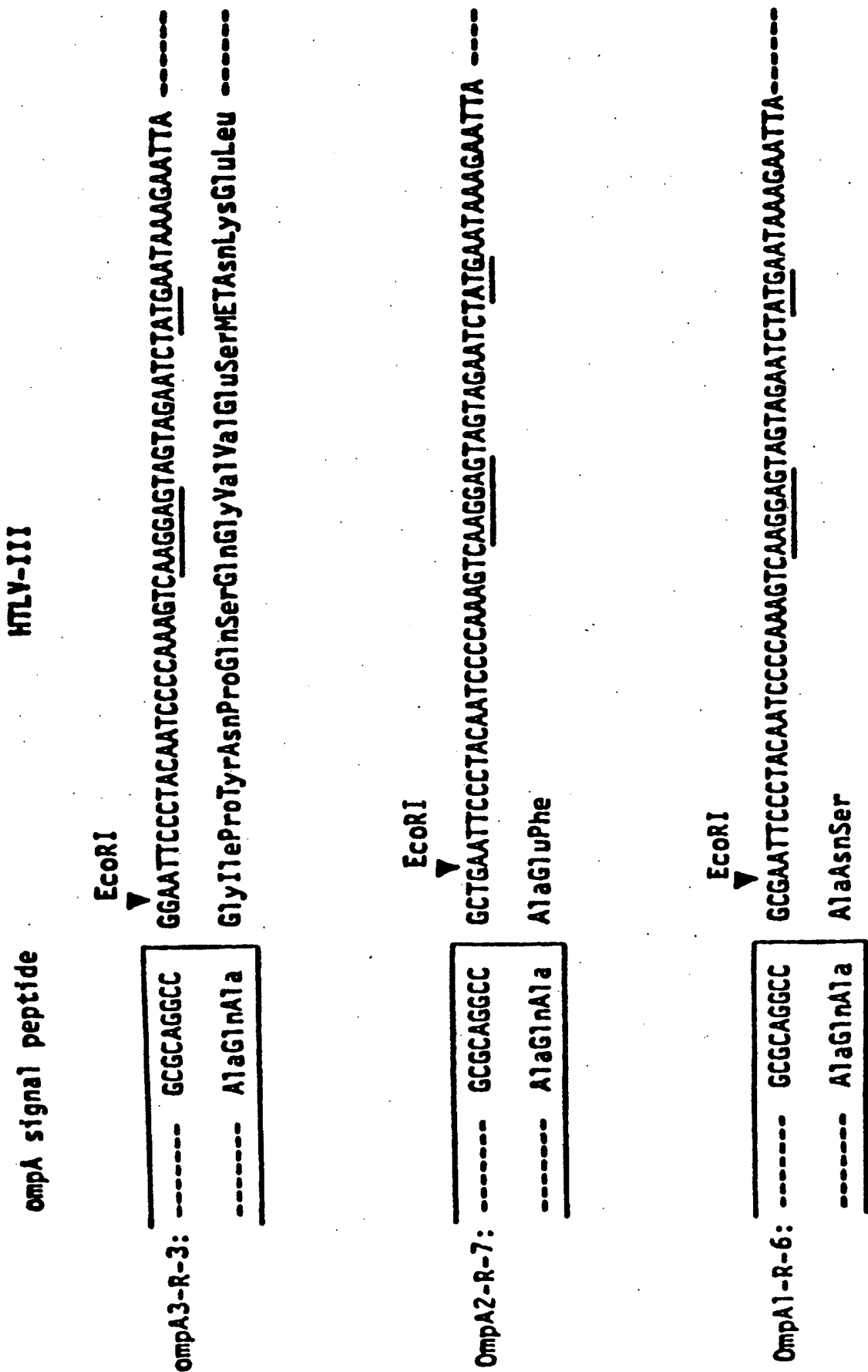


FIG. 6

